1-1-2014

Dynamic modeling of GNSS troposphere wet delay for estimation of precipitable water vapour

Ahmed El-Mowafy

Johnny Lo
Edith Cowan University, j.lo@ecu.edu.au

10.1515/jag-2013-0012
This is an Author's Accepted Manuscript of: El-Mowafy A., Lo J. (2014). Dynamic modeling of GNSS troposphere wet delay for estimation of precipitable water vapour. Journal of Applied Geodesy, 8(1), 31-42. Available here

Research article

Ahmed El-Mowafy* and Johnny Lo

Dynamic modeling of GNSS troposphere wet delay for estimation of Precipitable Water Vapour

Abstract: Proper dynamic modelling of the troposphere wet delay using the Global Navigation Satellite Systems (GNSS) measurements is important in precise point positioning and in estimation of the Precipitable Water Vapour (PWV) for weather forecast. The random walk (RW) and the first-order Gauss-Markov (GM) autocorrelation models are commonly used for this purpose. However, it was found that these models consistently underestimate the temporal correlations that exist among the troposphere wet delay. Therefore, a new dynamic model is proposed. The performance of the proposed model in following the autocorrelation of actual data is demonstrated and its impact on the near-real time estimation of the wet delay was tested and compared to that of the GM and RW models. Results showed that the proposed model outperformed these models. When the computed wet delays were used to compute PWV, their estimated values were very close to actual PWV data measured by radiosonde with differences less than 1 mm.

Keywords: GNSS, troposphere wet delay, dynamic modeling, Precipitable Water Vapour

***Corresponding Author: Ahmed El-Mowafy:** Department of Spatial Sciences, Curtin University, GPO BOX U 1987, Perth, WA 6845, Australia, E-mail: A.El-mowafy@curtin.edu.au
Johnny Lo: School of Engineering, Edith Cowan University Perth, WA, Australia

1 Introduction

The troposphere is the lower part of the atmosphere and extends from the Earth's surface up to an altitude of about 20 kilometers. In GNSS, the troposphere delay for all satellites observed at one position is traditionally modeled as one parameter projected along the zenith direction and a mapping function is applied to project it along each satellite-to-receiver direction. This delay can be divided into two components, the hydrostatic delay and the wet delay. The zenith hydrostatic delay (ZHD) can be estimated with empirical models Saastamoinen [17] to a few millimetres in accuracy. However, determination of the zenith wet delay (ZWD) represents a difficult task due to the dynamic nature of the atmospheric water vapour. Due to changes of the temporal and spatial variability of the water vapour, the wet delay cannot be consistently modeled with millimeter precision by any existing empirical model.

Precise estimation of the ZWD is essential for high-precision positioning applications, such as Network Real Time Kinematic (RTK) and in Precise Point Positioning (PPP). In addition, the ZWD values determined from GNSS measurements can be also used in Numerical Weather Prediction (NWP) modeling and to estimate the PWV. The use of these PWV derived from GNSS–ZWD and its impact on weather forecasting was discussed in [11, 22, 5, 23, 18, 15]. These studies reported improvements in the humidity and precipitation forecasts when GNSS PWV estimates are assimilated into NWP models.

Appropriate dynamic modeling of the troposphere is an important task to accurately estimate its value. The troposphere delay is linearly modeled as a bias in the GNSS observation model along with other biases (e.g. hardware biases, initial phase bias, etc.); hence, dynamic modeling can help in its distinct parameterization. In addition, when processing GNSS observations using Kalman filtering (KF), the use of a correct dynamic model is essential, otherwise the filter may diverge as the predicted troposphere delay through the dynamic model is used as a pseudo observation. In forming an appropriate model that describes dynamics of change of the tropospheric wet delay parameter with time, one needs to study the troposphere autocorrelation, which describes the temporal correlations between pairs of GNSS tropospheric estimates in a time series. The autocorrelation investigation also plays a role in determining the autocorrelation time length, which is an essential parameter needed in dynamic modeling.

In this study, different dynamic models of the troposphere and their performance are investigated with the purpose of recommending the best model. The paper is organized as follows. First, two of the more commonly-used dynamic models in the KF process, namely the RW and GM models, are outlined. A new dynamic model is then proposed to model the temporal transition of the ZWD. The proposed model is analyzed with real GNSS data, and its results are compared to those of the RW and GM models. The estimated wet troposphere using the developed model was used to estimate PWV, which was compared

with radiosonde reference values to assess the validity of the model.

2 Modelling of the troposphere in the GNSS observation equations

The GNSS code and phase observation equations can be formulated as [12]:

$$P(t) = \rho_i^k(t, t-\tau) + d\rho^k(t-\tau) + cdt_i(t) - cdt^k(t-\tau) + I_i^k + T_i^k + dp_i^k + IF + \varepsilon_i^k(\rho) \quad (1)$$

$$\phi(t) = \phi_i^k(t, t-\tau) + d\rho^k(t-\tau) + cdt_i(t) - cdt^k(t-\tau) - I_i^k + T_i^k + d\phi_i^k + N_i^k + \varepsilon_i^k(\phi) \quad (2)$$

where $P(t)$ and $\phi(t)$ are the code and the phase measurements received at time t, ρ_i^k is the receiver-to-satellite range, $d\rho^k$ is the orbital error, τ is the time taken by the signal to travel from the satellite to the receiver, c denotes the speed of light, dt_i and dt_k are the receiver i and satellite k clock errors. I_i^k is the ionosphere error, and T_i^k denotes the total troposphere delay. dp_i^k is the receiver and satellite hardware code biases and $d\phi_i^k$ includes the receiver and satellite hardware phase biases and the initial phase biases. Both terms also include smaller errors such as the relativistic error, Sagnac delay, receiver and satellite antenna-phase centre offsets and variations, site displacement effects due to Earth tide, ocean tide and atmospheric loading [2]. IF denotes the inter-frequency bias, and N_i^k is the integer phase ambiguity. Finally, $\varepsilon_i^k(\rho)$ and $\varepsilon_i^k(\theta)$ are the code and phase noises, which are usually assumed Gaussian with zero mean. Most of the errors are minimized by differencing over short to medium distances, and in case of the ionosphere, its first order term can be eliminated by the use of dual-frequency ionosphere-free linear combination of observations.

The model of equations (1) and (2) is rank deficient if the slant troposphere error for each satellite is to be estimated using least squares adjustment. To minimize rank deficiency, the troposphere delay is generally expressed at each ground location in terms of one value taken along the zenith, i.e. Zenith Total Delay (ZTD). This one value of ZTD is used for all satellites observed from a single location where a mapping function is applied to project the ZTD onto the receiver-to-satellite line of sight direction for each satellite, such that [9]:

$$T_i^k - m(\theta_i^k)ZTD_i \quad (3)$$

where θ_i^k is the elevation angle between the receiver i and the satellite k, $m(\theta_i^k)$ is the mapping function and ZTD_i is the ZTD at receiver i. Traditionally, the hydrostatic and wet components of the total troposphere delay are treated separately. In this case, two mapping functions are needed for the hydrostatic and wet delays, denoted as m_H and m_W. The total troposphere delay can then be expressed as:

$$T_i^k = m_H(\theta_i^k)ZHD_i + m_W(\theta_i^k)ZWD_i \quad (4)$$

where ZHD_i and ZWD_i are the Zenith hydrostatic and wet components of the total troposphere delay at station i. When surface pressure is available, the ZWD can be the extracted from the estimated ZTD value by accounting for the ZHD using empirical models such as the Saastamoinen [17] or Hopfield [8] dry models.

In practice, a single value for the ZWD parameter is generally estimated for a one to two hour interval [10]. This is due to the fact that the ZWDs generally do not vary significantly from their mean value during these short time intervals, i.e. the ZWD data behaves like a stationary process. As an example, Figure 1 provides an illustration of the ZWD variation around its mean for a two hour period. The ZWD data in this figure is estimated from Water Vapour Radiometer (WVR) observations at the International GNSS Service (IGS) station Onsala on the 10th of September in 2003. By assuming a constant mean value, $\overline{ZWD}$, over a short time-period, the ZWD can be given as:

$$ZWD_i = \overline{ZWD} + \Delta ZWD_i \quad (5)$$

where ΔZWD_i is the difference between the ZWD value at time i and the mean value $\overline{ZWD}$. Another approach is to estimate a rough estimate of the mean parameter $\overline{ZWD}$ via empirical wet delay models [4].

If Kalman filtering is used to process the data, the rank deficiency due to the slant troposphere for each satellite will not be present due to the use of dynamic modeling (time update) as each predicted unknown including the slant troposphere is treated as a pseudo observation. Nevertheless, most practitioners estimate the troposphere as a single value along the zenith in Kalman filtering and apply a mapping function, primarily to simplify the computations. At the initial epoch, the state vector X_0, which includes the troposphere and its covariance matrix Q_{X_0} are typically assumed to be known. The state dynamic model that relates two consecutive values of the state vector, i.e. X_i and X_{i-1} at times i and $i-1$, reads:

$$X_i = \Phi_{i/i-1}X_{i-1} + u_i \quad (6)$$

where $\Phi_{i/i-1}$ is the transition matrix. The system noise u_i is assumed to follow a normal distribution with zero mean and a known covariance matrix Q_u. Using the covariance

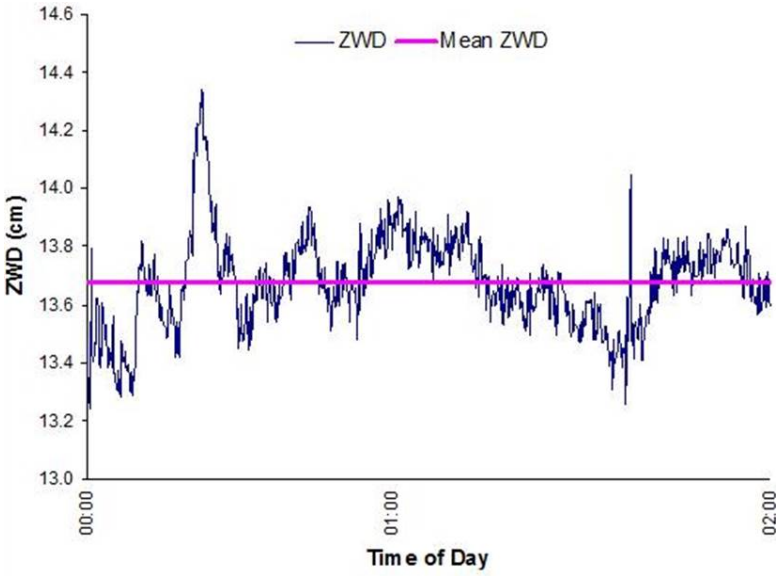

Fig. 1. A plot of the time series of WVR ZWD data at the Onsala station in a 2 hr period.

propagation law, the covariance matrix of the predicted state vector is given by [24]:

$$Q_{\hat{X}_i/i-1} = \Phi_{i/i-1} Q_{\hat{X}_{i-1/i-1}} \Phi^T_{i/i-1} + Q_u \tag{7}$$

The dynamic model of the wet delay is not only needed to reflect changes of the user position in the spatial domain. In static positioning, when the troposphere error is estimated at a known point, the positional state vector, X_i does not change with time, i.e. $X_i = X_{i-1}$, the corresponding $\Phi_{i/i-1} = I$, and Q_u is assumed to be zero. However, assuming estimating the ZHD parameter using one of the empirical methods, the ZWD parameter will vary with time due to the fluctuations of the water vapour in the atmosphere [14]. In this instance, an appropriate representation of the transition between adjacent ZWD measurements is needed.

The next sections will focus on the dynamic modeling of the ZWD through addressing the transition matrix, estimation of its parameters and the corresponding stochastic parameters of Q_u.

3 Autocorrelation models for ZWD

In this section, different widely used dynamic models will first be presented and their drawback will be discussed. A new dynamic model is proposed that can overcome these drawbacks.

3.1 Traditional Dynamic models for ZWD

A random walk (RW) model can be used for dynamic modeling of the ZWD. It defines a random process whereby the value of the ZWD_i is composed of the past variable ZWD_{i-1} plus an error term defined as a white noise ε_i with zero mean such that:

$$ZWD_i = ZWD_{i-1} + \varepsilon_i \tag{8}$$

The associated variance of the RW process noise ε_i is:

$$E(\varepsilon_i^2) = \rho^2 \Delta t^2 \tag{9}$$

where ρ^2 is the variance of the RW process and Δt is the time interval. A drawback of this model is that the variance of the RW process noise always grows with time.

The first-order Gauss Markov (GM) model can also be used to describe the temporal changes of the ZWD assuming that the correlations among the ZWD decays smoothly with time, such that the dynamic model reads:

$$ZWD_i = e^{-\frac{1}{\tau_{GM}}\Delta t} + u_i \tag{10}$$

where τ_{GM} is the correlation time of the GM model, and u_i is a white noise with zero mean and covariance Q_u. The associated variance of the GM process noise is given as:

$$E(u_i^2) = \sigma_{GM}^2 \left(1 - e^{-\frac{2}{\tau_{GM}}\Delta}\right) \tag{11}$$

where Δ_{GM}^2 is the steady-state variance of the GM process. Using equation (5), the GM model given by equation (10) can then be expressed as:

$$\begin{aligned} ZWD_i &= \overline{ZWD} + e^{-\frac{1}{\tau_{GM}}} \Delta ZWD_{i-1} + \tilde{u}_i \\ &= \overline{ZWD} + \Theta_{i,i-1} \Delta ZWD_{i-1} + \tilde{u}_i \end{aligned} \tag{12}$$

where $\tilde{u}_i$ is a white noise for ΔZWD with zero mean and variance $\sigma_{\tilde{u}}^2$. This variance is identical to that given by equation (11).

The GM correlation time can be estimated from the autocorrelation function $\{\rho(\Delta t) = e^{-\frac{1}{\tau_{GM}}}\}$ at the point $\rho(\Delta t) = \frac{1}{e}$ when $\tau_{GM} + \Delta t$. Alternatively, it can be determined at a specific time lag where significant ZWD autocorrelation is no longer observed. For instance, Figure 2 shows the autocorrelation of PWV (estimated from the ZWD) with lags of 1 hr intervals at ALIC station in Australia at three different dates 31 March 2010 (Figure 2A), 3rd April 2010 (Figure 2B), and 6th April 2010 (Figure 2C). From the figure, τ_{GM} can be determined by finding the intersection between the autocorrelation trend line and the confidence interval [3], which varies within a small range between different tests. From the figures, the value of τ_{GM} can be taken between 1 and 2 hours. A drawback of the GM model is that it overestimates the temporal decrease rate of ZWD as will be shown through an example in the next section.

3.2 A Proposed Autocorrelation Model

An alternative autocorrelation function for ZWD is proposed to be used in the transition matrix of the dynamic model. Analytical studying of the autocorrelation of ZWD of several data sets shows that the trend exhibited by a hyperbolic function gives a reasonable representation of this autocorrelation changes. Thus, the proposed autocorrelation function between the ZWDs at epochs i and $i - \Delta t$, i.e. for a time lagΔt, can be given by:

$$\rho(\Delta t) = \frac{1}{\left(\frac{\Delta t}{\tau_{PM}} + 1\right)^{\frac{\Delta t}{\tau_{PM}\beta}}} \tag{13}$$

and the state element of the ZWD at time i can be represented by:

$$ZWD_i = \frac{1}{\left(\frac{\Delta t}{\tau_{PM}} + 1\right)^{\frac{\Delta t}{\tau_{PM}\beta}}} \times ZWD_{i-1} + u_i \tag{14}$$

where τ_{PM} is the correlation time of the proposed model, and the parameter β is either chosen based on the analysis of several previous data sets or to be determined from an initial period of the data at hand. For instance, for a set of n autocorrelation estimates that is determined using a standard autocorrelation approach, for example:

$$\rho(\Delta t) = \frac{\hat{Z}(t_i + \Delta t)}{\hat{Z}(t_i)} \tag{15}$$

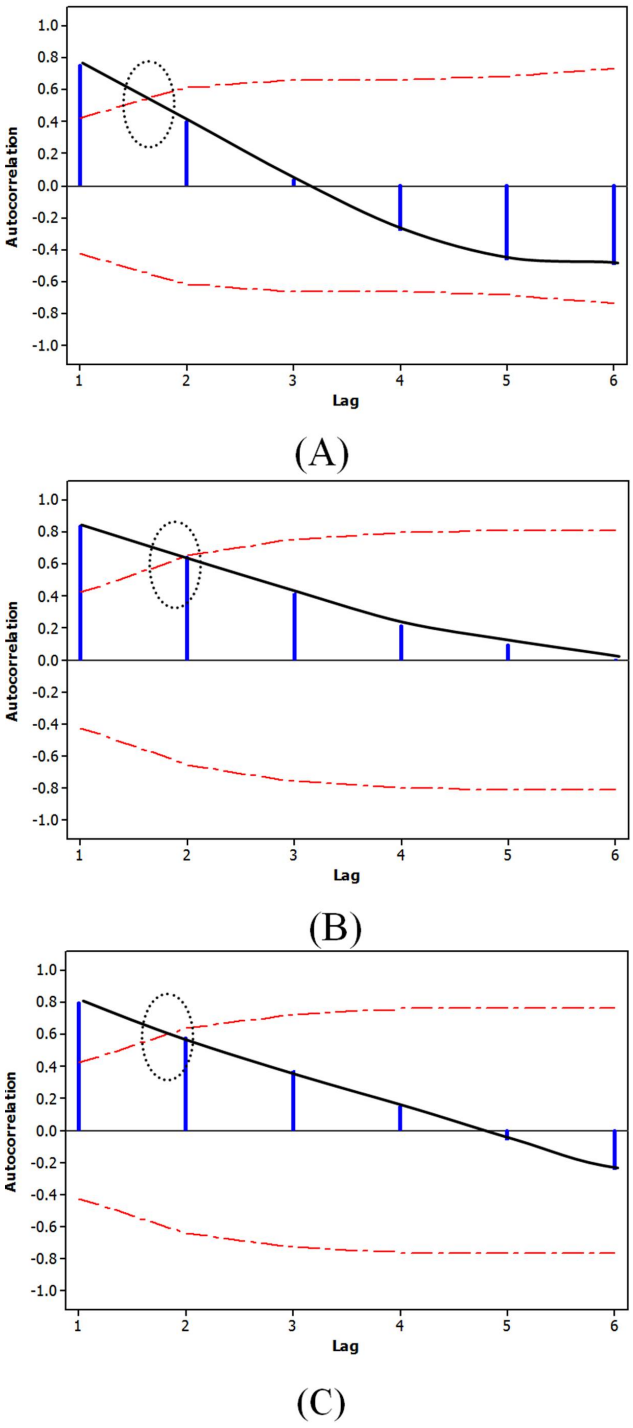

Fig. 2. Autocorrelation plot of the PWV estimates over ALIC and its confidence interval w.r.t the time lag in hrs showing when the autocorrelation becomes insignificant.

with

$$\hat{Z}(t_i + \Delta t) = \frac{\sum_{i=1}^{n-1}(ZWD(t_i) - \overline{ZWD}(t_i + \Delta t) - \overline{ZWD})}{n} \quad (16)$$

Taking the natural logarithm of both sides of equation (13) results in the linearized form:

$$\ln(\rho) = \left[-\left(\frac{\Delta t}{\tau_{PM}}\right) \times x \ln\left(\frac{\Delta t}{\tau_{PM}} + 1\right)\right]\beta \quad (17)$$

An estimate for β i.e. $\hat{\beta}$ can then be calculated by performing least-squares analysis on the first $\frac{n}{12}$ number of autocorrelation values (as traditionally applied), generated by equations (13) and (15), using the linear relationship defined by equation (17). In real-time applications, a default value of β can be used during this period until $\hat{\beta}$ is computed [4]. Once $\hat{\beta}$ has been determined, the proposed model given by equation (14) is then fully defined. To determine the variance of the process noise, the noise u_i is once again isolated, then taking the expectation of its square with σ^2, which is the variance of the process, gives:

$$\begin{aligned} E(u_i^2) =& E\left(ZWD_i - \frac{1}{\left(\frac{\Delta t}{\tau_{PM}} + 1\right)^{\frac{\Delta t}{\tau_{PM}\hat{\beta}}}} \times ZWD_{i-1}\right)^2 \\ =& \sigma^2\left(1 - \left(\frac{1}{\left(\frac{\Delta t}{\tau_{PM}} + 1\right)^{\frac{2\Delta t}{\tau_{PM}}}}\right)\right) \end{aligned} \quad (18)$$

To evaluate the proposed model, Figure 3 to Figure 6 demonstrate the capability of this model in following the trend of autocorrelations, calculated using equation (15), of actual PWV values (as representative of ZWD). The PWV were determined from radiosonde data collected for three days at four different locations across Australia (Alice Springs, Broome, Burnie and Ceduna). The GM model is also included in these figures for comparison purposes. For the GM model, the value of τ_{GM} is determined at a time lag Δt where statistically significant autocorrelation is observed using the Ljung-Box Q statistic [13]. For the proposed model, τ_{PM} is taken equals τ_{GM}.

From the Figure 3 to Figure 6, it can be seen that the GM autocorrelation function did not adequately represent the actual PWV autocorrelations. It consistently overestimated the rate at which the PWV autocorrelation values decreases. Conversely, the proposed model was able to provide autocorrelation that closely follow the actual values for a significant length of time.

4 Performance of the Proposed Model in Near Real-time Estimation of ZWD

In this section, the impact of the proposed model (PM) on the near real-time estimation of the ZWD is tested. The corresponding results are compared to that of the GM and RW models as these models are the current widely used models for ZWD estimation. 24 hours of GNSS dual-frequency data with 30 seconds sample intervals on the 25th Jan 2010 from two Western Australian IGS stations, namely Yarragadee (YAR2) and Karratha (KARR), were used to test the models. The stations were processed independently in the PPP mode. IGS products, including the IGS final orbital file, satellite clock information, Earth Orientation Parameters (EOPs), the coordinates of the ground stations and the antenna phase centre offsets and variations were used in the PPP processing [2]. An elevation angle cut-off of 5° and the Niell mapping functions [16] were used. The ionosphere-free linear combination of GNSS observations was implemented to mitigate the first-order ionosphere residual errors.

In conjunction with the surface meteorological data (humidity and pressure), the Saastamoinen hydrostatic model was used to provide *a-priori ZHD* estimates, which has an approximate accuracy of 95% [21]. These ZHD estimates, with the aid of the dry mapping functions were then subtracted from the observations leaving mainly behind the ZWD parameters in the troposphere term to be estimated. KF was used to estimate the ZWD along with the station coordinate, ambiguities and receiver clock error. The RW, GM and proposed models were used for dynamic modeling of ZWD in three separate runs of KF. The station coordinates were not assumed fixed as this part of our study is carried out to mimic kinematic positioning. For the proposed model, two approaches were used:

- PM1, where the ZWD is estimated as a random process in the form:

$$ZWD_i = \frac{1}{\left(\frac{\Delta t}{\tau_{PM}} + 1\right)^{\frac{\Delta t}{\tau_{PM}\hat{\beta}}}} ZWD_{i-1} + u_i$$

- PM2, where the ZWD is estimated in terms of the mean $\overline{ZWD}$ and the residual component simultaneously, i.e. in the form:

$$ZWD_i = \overline{ZWD} + \frac{1}{\left(\frac{\Delta t}{\tau_{PM}}\right)^{\frac{\Delta t}{\tau_{PM}}\beta}} \Delta ZWD_{i-1} + \tilde{u}_i$$

An autocorrelation analysis of the ZWD estimates across 10 Australian stations was first carried out. Based on the

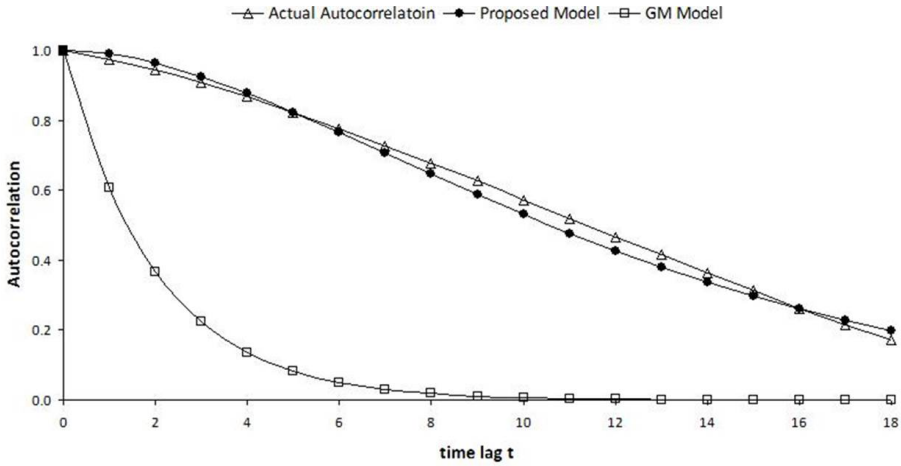

Fig. 3. Comparison among autocorrelations of the proposed and the GM models with the actual PWV at ALIC showing the divergence of GM and the ability of the proposed model to closely trace the actual PWV.

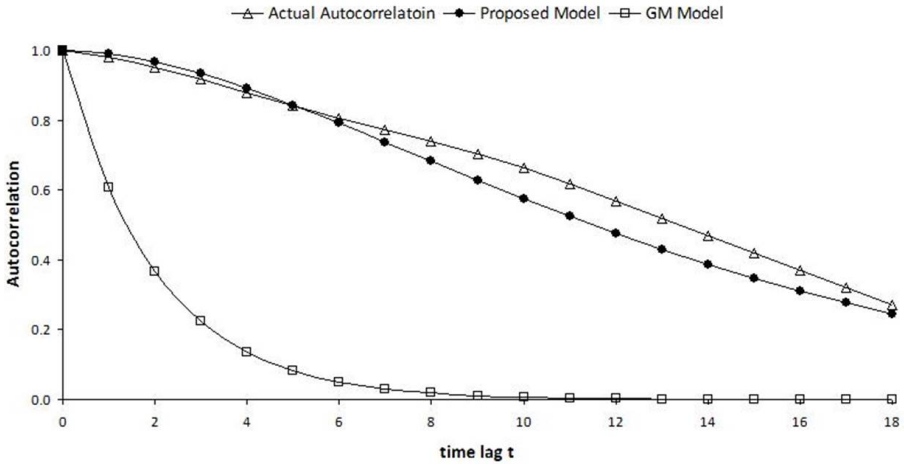

Fig. 4. Comparison between the proposed model and the GM model in estimating the actual PWV autocorrelations at Burnie (Tasmania)- the proposed model again well represent the autocorrelation of the actual PWV.

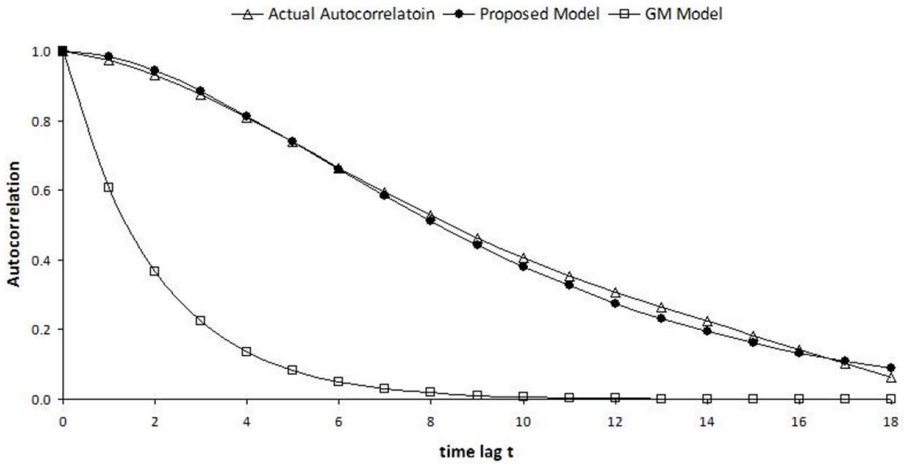

Fig. 5. Comparison between the proposed model and the GM model in estimating the actual PWV autocorrelations at Burnie - the proposed model best represent the autocorrelation of the actual PWV.

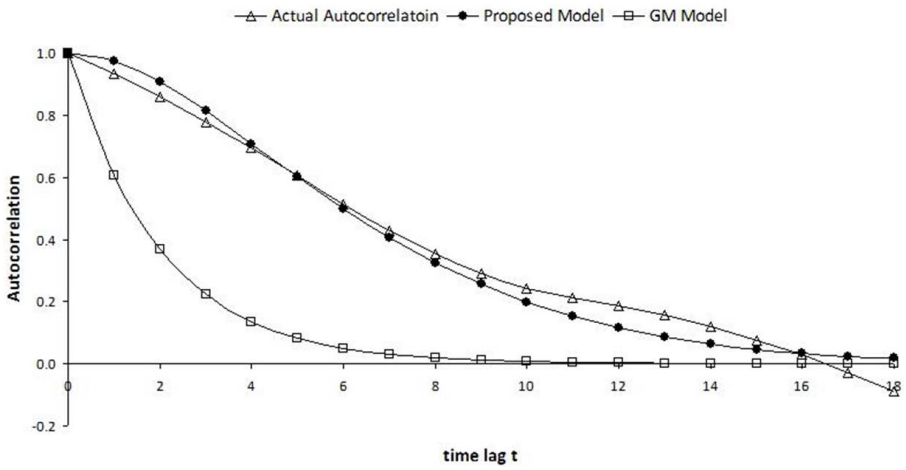

Fig. 6. Comparison between the proposed model and the GM model in estimating the actual PWV autocorrelations at CEDU - the proposed model best represent the autocorrelation of the actual PWV.

analysis of the ZWD autocorrelation results, the correlation time for both the GM model and the proposed model was empirically taken as 4800 seconds. The empirical β value for the PM is estimated as 0.75.

Once the ZWD is estimated from PPP processing of the data in our test sites (YAR2 and KARR), it is then added to the estimated ZHD to compute the ZTD. The estimated ZTD from each of the models are averaged at every 5 min and at every two hour periods, during the course of the 24 hr test period and are then compared to two sets of the IGS troposphere delay products that are sampled at 5 min and 2 hr intervals, which were taken as the reference for our comparison. Table 1 presents the Root Mean Square Error (RMSE) computed from the differences between the estimated ZTD and the 5 minutes IGS ZTD solution, whilst Table 2 provides the RMSE of the estimated ZTD when differenced from the 2 hr IGS solution. The tables include results when assuming 2 mm, 5 mm, and 10 mm standard deviation of the process noise for the ZWD.

Table 1 and Table 2 show that, in general, the RW model was the worst performer. The PM1, generally produced the best results at station KARR and comparable results to the GM model at station YAR2, with the corresponding ZTD RMSE values ranging from 10 mm to 20 mm. In most of the cases, PM1 gave better results compared to PM2. Overall, the best results were achieved at standard deviation of the process noise of 5 mm. There were marginal RMSE differences when the estimated ZTD were compared to the 5 min and the 2 hr IGS solutions across both test stations. Taking RMSE1 as the RMSE values when the estimated ZTD were referenced to the 5 min sampling rate solution, and RMSE2 as the RMSE values when they were referenced the 2 hr sampling rate solution. The maximum difference between RMSE1 and RMSE2 was 2.4 mm, with an average difference of 0.7 mm. This indicates that the difference between the two solutions is practically not significant. Figure 7 provides a plot of the spread of the differences between RMSE1 and RMSE2.

A difference of a few mm can be observed between the PM1 and PM2 ZWD estimates. This can be explained by examing the parameterisation of $\overline{ZWD}$ and ΔZWD in the corresponding design matrix in KF. In modelling the $\overline{ZWD}$, the coefficients of its corresponding column in the design matrix is a vector of ones. The coefficients for the ΔZWSs which were modelled by the PM, are close to one due to the high correlation between successive ZWD estimates in this test since the sampling interval was only 30 seconds. The design matrix will therefore include two columns that are almost similar. Thus, to avoid singularity in this case, the use of the PM2, where the ZWD is decomposed into the two components ZWD and ΔZWD, is recommended only when processing longer time intervals or when $\overline{ZWD}$ is estimated in advance and is reduced from the observations.

5 Accuracy of GNSS-ZWD for Estimation of PWV Using the Proposed model

The general consensus from past studies [1, 6, 19, 20] is that a set of GNSS ZWD estimates for the determination of PWV is considered as good estimates if it has RMSE values of less than 15-20 mm in ZWD (giving 2-3 mm in PWV) when referenced to RS estimates, or any other reliable independent measurements such as WVR or Very Long Base-

Table 1. RMSE (mm) of the differences between the estimated ZTD and the IGS solutions (5 min sampling rate).

			1 mm SD	
Station	RW	GM	PM1	PM2
YAR2	17.2	16.1	14.2	14.3
KARR	17.1	28.6	21.8	20.5
			5 mm SD	
Station	RW	GM	PM1	PM2
YAR2	15.3	14.6	14.7	14.4
KARR	11.4	11.1	10.1	11.7
			10 mm SD	
Station	RW	GM	PM1	PM2
YAR2	15.6	14.4	15.1	15.8
KARR	14.9	13.9	13.5	15.8

Table 2. RMSE (mm) of the differences between the estimated ZTD and the IGS solutions (2 hr sampling rate).

			1 mm SD	
Station	RW	GM	PM1	PM2
YAR2	15.8	15.2	14.6	12.2
KARR	15.7	26.5	20.7	18.9
			5 mm SD	
Station	RW	GM	PM1	PM2
YAR2	14.0	13.3	13.4	13.6
KARR	12.2	10.9	10.3	11.5
			10 mm SD	
Station	RW	GM	PM1	PM2
YAR2	15.2	14.3	14.7	15.6
KARR	15.9	14.9	14.2	15.3

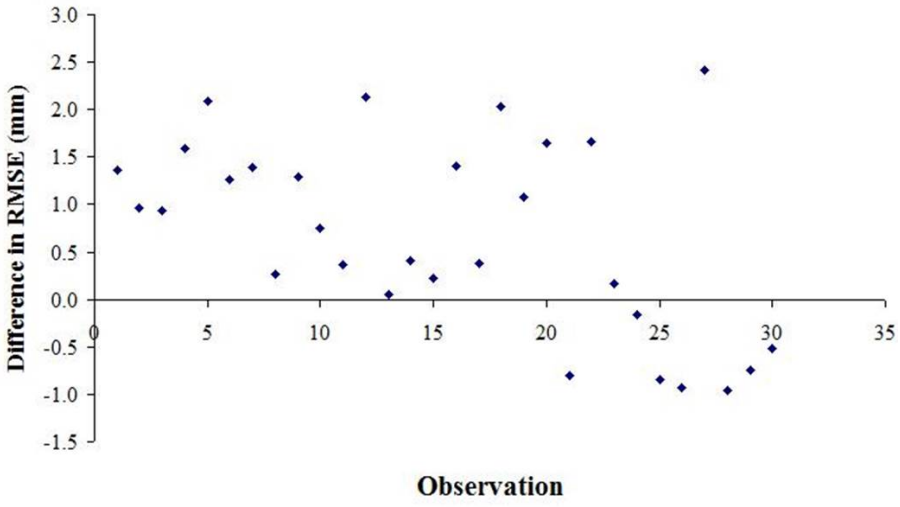

Fig. 7. Differences between RMSE1 (ZTD referenced to the 5 min IGS sampling rate solution) and RMSE2 (ZTD referenced to the 2 hr IGS sampling rate solution).

line Interferometry (VLBI) solutions. For assimilation purposes, it is preferable that the accuracy of the GNSS ZWD estimates is within 7-15 mm, or 1-2 mm in PWV [15].

To assess accuracy of ZWD values estimated by using the proposed model, they should first be validated. The process of data validation and evaluation of accuracy can be performed at locations where RS, WVR or VLBI reference values are available. To validate GNSS ZWD results, one may set up an acceptance null hypothesis, assuming that the GNSS ZWD deemed adequate (in a statistical sense) if:

$$H_0 : ZWD_{GNSS} \approx ZWD_{RS} \tag{19}$$

and the alternative hypothesis is that they are not equal, where ZWD_{GNSS} and ZWD_{RS} denote the ZWD values estimated from GNSS and the reference system (e.g. RS), respectively. A ZWD value can be considered as an *outlier* if:

$$|ZWD_{GNSS} - ZWD_{RS}| \geq t_{\alpha_s/2,n-1} \times s_{ZWD_{GNSS}} \tag{20}$$

where $s_{ZWD_{GNSS}}$ is the standard error of the GNSS ZWD estimate computed from the covariance matrix and the sample size. $t_{\alpha_s/2,n-1}$ denotes the upper $\alpha_s/2$ percentage point of the t–distribution with α_s significance level. It is assumed here that the discrepancies between ZWD_{GNSS} and ZWD_{RS} come from a population that is normally distributed, and that prior information regarding the population true variance σ^2 is unknown. The P-value, which is the probability of observing a sample statistic as extreme as the test statistic, was computed using the t–distribution for the value ($t_{obs} = |ZWD_{GNSS} - ZWD_{RS}|/s_{ZWD_{GNSS}}$).

Assuming a significance level of 0.05, the P-value is compared to the significance level. When the P-value is larger than the significance level; Ho is not rejected, and the test concludes that there is *no significant statistical difference* between the GNSS and RS ZWD estimates, and indicating that the error estimate provides a realistic measure of the quality of the ZWD solution. This error information can then be used to weight the ZWD observations in the NWP assimilation process.

If an outlier in ZWD_{GNSS} is detected, an investigation into the cause of such outlier should be carried out. Too many outliers may indicate the existence of a bias and corrective action such as re-sampling of the data points may be necessary (Montgomery, 2001). It may also be necessary to obtain a more reliable mean estimate by increasing the sampling window size.

To assess the accuracy of PWV computed from GNSS estimated ZWD [6], they were compared with reference RS PWV in an independent test in a static mode. The ZWD_{GNSS} were determined using the proposed dynamic model with parameters determined from the test given in Section 3 and validated using the above methodology. The test was performed at five reference stations in Australia of different climatic regions, including the tropical north (station TOW2), the Mediterranean of the south-west (stations YAR2 and CEDU), the humid and cool subtropical of the east (station SHEP), and the arid center of Australia (station ALIC). The RS PWV daily data were available from nearby RS launch sites provided by The Australian Bureau of Meteorology. The approximate distance between the GNSS stations and radiosonde sites (in km) as well as the number of RS per day at each site are given in Table 3. In processing of the GNSS measurements, the positional information was assumed fixed as the test site has a known position. The remaining parameters, including phase ambiguities, clock errors etc., are estimated or modeled out of the observation equation beforehand. The ZHD was determined via the Saastamoinen hydrostatic model and subtracted from the ZTD parameter to estimate ZWD.

The PWV estimated from the ZWD_{GNSS} hourly values closest to the RS sample time were used in this comparison. The test spans 22 days (31 March to 21 April). This period has high diurnal variation and it allows GNSS to demonstrate its capability under varying atmospheric conditions. Table 3 shows that the $RMSE$ of the GNSS PWV when referenced to the RS PWV was in general less than 2 mm for short-medium separation distance between their sites, and due to spatial decorrelation the $RMSE$ increases with the increase in this distance. As an example, Figure 8 illustrates PWV computed from the GNSS ZWD, PWV from RS and their differences at station TOW2. The figure shows that the PWV difference (assumed as an error in the GNSS-derived PWV) was in general less than 1 mm (note the right vertical scale of the figure). The differences in PWV appears to have a bias, which can be attributed to three factors; the spatial separation between the locations of the RS and GNSS data collection sites; accuracy of estimation of the empirical ZHD; and the time shift between the closest hourly GNSS-derived PWV with the daily RS data. Figure 9 illustrates a regression plot between the GNSS and RS PWV estimates, where a regression correlation R_{reg} value of $\sqrt{0.8873} = 0.9420$ was observed, which indicates a strong agreement between the GNSS-estimated PWV and RS-measured PWV.

6 Summary and Conclusions

The first-order Gauss-Markov (GM) autoregressive function is widely used for modeling the dynamic behaviour of the ZWD. To investigate its performance in modeling actual

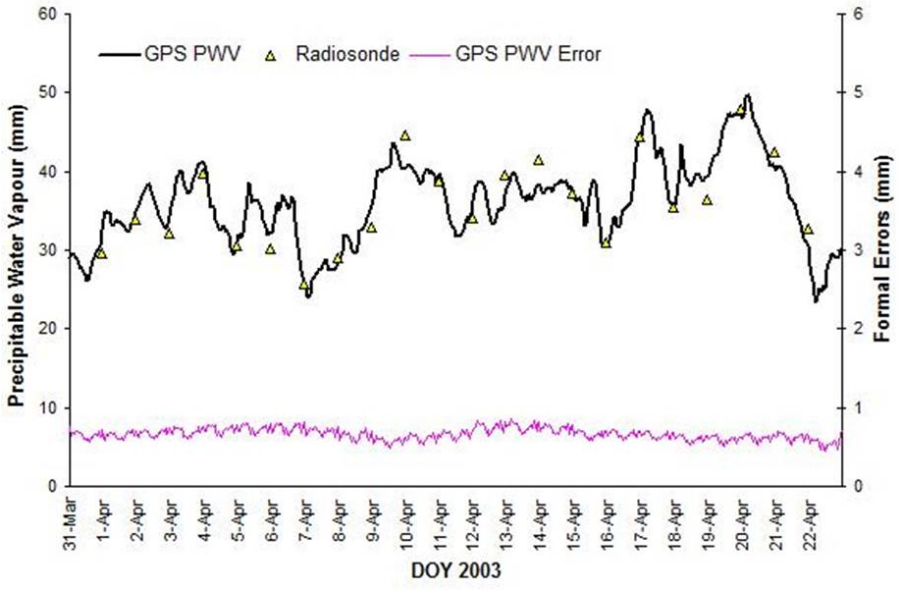

Fig. 8. Time series of the GNSS (using the proposed model) and RS PWV estimates and their differences at station TOW2.

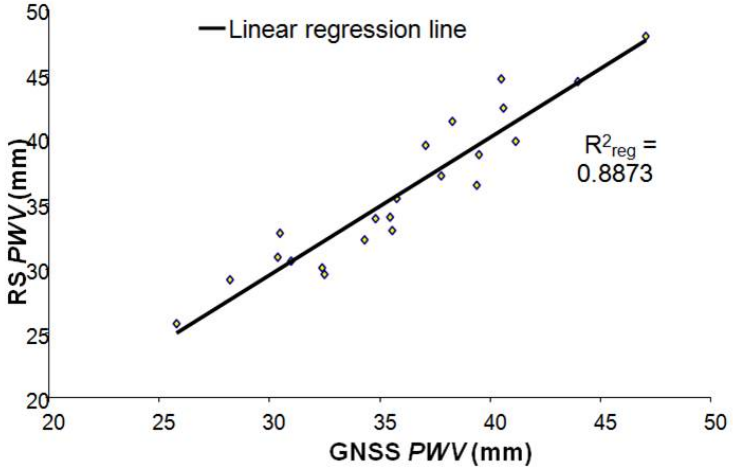

Fig. 9. Regression plot between the GNSS and RS PWV estimates at TOW2.

Table 3. RMSE (mm) of the GNSS-RS PWV.

Station	distance between the GNSS and radiosonde sites (km)	Number of RS per day	PWV RMSE (mm)
ALIC	14	1	1.04
TOW2	30	1	0.95
YAR2	69	1	2.10
SHEP	153	2	3.38
CEDU	293	2	4.62

PWV data (as representative of *ZWD*), radiosonde *PWV* data were collected at four locations across Australia and the trends of their computed autocorrelations were compared with autocorrelations from the GM model. It was found that the GM model consistently underestimates the temporal correlations of the *PWV* measurements. Therefore, a new autocorrelation dynamic model is proposed. The proposed model gave results in good agreement with the autocorrelation changes of the actual *PWV* for the test data considered.

The impact of the proposed dynamic model on the near-real time estimation of the *ZWD* was also tested and its results were compared to that of the GM model as well as the random walk model. In this test, 24 hours of GPS dual-frequency data collected at two Western Australian IGS stations were used. The data were processed independently in a PPP mode using each of the three models. The published IGS final *ZTD* at the two stations were used as a reference for comparison of the results. In estimation of the *ZWD*, two approaches were considered. The first is a classical approach where *ZWD* is modeled as one variable. In the second approach, the *ZWD* is considered as comprising two components, a mean value that is taken constant over short time-periods, and a variable component that is modeled as a random process. Results at the two stations showed that the proposed autocorrelation model generally produced the best results, with the corresponding *ZTD* RMSE values ranging from 10 mm to 20 mm. The traditional approach, where *ZWD* is estimated as one value, and the approach where the *ZWD* is estimated as a mean value and a random process gave comparable results.

The accuracy of the *PWV* values computed from GNSS estimated *ZWD* using the proposed dynamic model was assessed by comparing them with actual data measured by a radiosonde. The test data spans 22 days. A strong agreement was observed between the GNSS estimated *PWV* and the actual *PWV* with differences less than 1 mm when their locations are separated by relatively short distances.

Acknowledgments

This work is supported by an Australian Research Council APAI grant No. LP0347723, the Australian Bureau of Meteorology, and a PhD scholarship to the second author from the Department of Spatial Sciences, Curtin University, Perth, Australia. Dr. C. Hu is acknowledged for processing some of the data using PPP software.

References

[1] Businger S., Chiswell S. R., Bevis M., Duan J., Anthes R. A., Rocken C., Ware R. H., Exner M., Van Hove T. and Solheim F. S., The promise GPS in atmospheric monitoring, *Bulletin of the American Meteorological Society* **77**(1) (1996), 5-18.

[2] El-Mowafy, A., An Alternative Post-Processing Relative Positioning Approach Based on Precise Point Positioning, *Journal of Surveying Engineering*, ASCE **135**(2) (2009), 56-65.

[3] El-Mowafy, A. and Lo. J., Prediction of Troposphere Wet Delay, *Journal of Applied Geodesy* **5**(3-4) (2011), 163-173.

[4] El-Mowafy A., Lo. J. and Hu C., Dynamic Modelling of Zenith Wet Delay in GNSS Measurements, *Proceedings of ION International Technical Meeting*, San Diego, CA, 24-26 Jan, 2011, 965-972.

[5] Gutman, S. I., Sahm S. R., Benjamin S. G., Schwartz B. E., Holub K. L., Stewart J. Q. and Smith T. L., Rapid retrieval and assimilation of ground based GPS precipitable water observations at the NOAA forecast systems laboratory: Impact on weather forecasts, *Journal of the Meteorological Society of Japan*, **82**(1B) (2004), 351-360.

[6] Haase J. S., Ge M., Vedel H. and Calais E., Accuracy and variability of GPS tropospheric delay measurements of water vapor in the western Mediterranean, *Journal of Applied Meteorology* **42**(11) (2003), 1547-1568.

[7] Heinkelmann R., Boehm J., Schuh H., Bolotin S., Engelhardt G., MacMillan D. S., Negusini M., Skurikhina E., Tesmer V. and Titov O., Combination of long time-series of troposphere zenith delays observed by VLBI, *Journal of Geodesy* **81**(6-8) (2007), 483-501.

[8] Hopfield H., Two-Quartic Tropospheric Refractivity Profile for Correcting Satellite Data, *Journal of Geophysical Research Oceans* **74**(18) (1969), 4487-4499.

[9] Kouba J., Implementation and testing of the gridded Vienna Mapping Function 1 (VMF1), *Journal of Geodesy* **82**(4,5) (2008), 193-205.

[10] Kouba J., A guide to using International GNSS Service (IGS) Products. Accessed 23/05/2013 at http://acc.igs.org/UsingIGSProductsVer21.pdf.

[11] Kuo Y. H., Zou X. and Guo Y. R., Variational assimilation of precipitable water using a nonhydrostatic mesoscale adjoint model, *Monthly Weather Review* **124**(1) (1996), 122-147.

[12] Leick A., *GPS Satellite Surveying*, 3rd ed. New Jersey, USA: John Wiley & Sons, Inc., 2004.

[13] Ljung G. M. and Box G. E. P., On a measure of a lack of fit in time series models, *Biometrika* **65** (1978), 297-303.

[14] Lo. J. and El-Mowafy A., Investigation into the Correlations among GNSS observations and their Impact on Height and Zenith Wet Delay Estimation for Medium and Long Baselines, *Geo-Spatial Information Science* **15**(4) (2012), 219-228.

[15] Macpherson S. R., Deblonde G., Aparicio J. M. and Casati B., Impact of NOAA ground-based GPS observations on the Canadian regional analysis and forecast system, *Monthly Weather Review* **136**(7) (2007), 2727-2745.

[16] Niell A. E., Global mapping functions for the atmosphere delay at radio wavelenghts, *Journal of Geophysical Research* **101**(B2) (1996), 3227-3246.

[17] Saastamoinen J., Contributions to the theory of atmospheric refraction, *Bullétin Géodésique* **105-107** 279-298, 383-397, (1973), 13-34.

[18] Smith T. L., Benjamin S. G., Gutman S. I. and Sahm S., Short-range forecast impact from assimilation of GPS-IPW observations into the Rapid Update Cycle, *Monthly Weather Review* **135**(8) (2006), 2914-2930.
[19] Snajdrova K., Boehm J., Willis P., Haas R. and Schuh H., Multi-technique comparison of tropospheric zenith delays derived during the CONT02 campaign, *Journal of Geodesy* **79**(10-11) (2006), 613-623.
[20] Steigenberger P., Tesmer V., Krügel M., Thaller D., Schmid R., Vey S. and Rothacher M., Comparisons of homogeneously reprocessed GPS and VLBI long time-series of troposphere zenith delays and gradients, *Journal of Geodesy* **81**(6-8) (2007), 503-514.
[21] Tuka A. and El-Mowafy A., Performance Evaluation of Different Troposphere Delay Models and Mapping Functions, *Measurement* **46**(2) (2013) 928-937.
[22] Vedel H. and Huang X. Y., A NWP impact study with ground based GPS data, *The International Workshop on GPS Meteorology*, Tsubaka, Japan, 2003.
[23] Vedel H. and Huang X. Y., Impact of ground-based GPS data on numerical weather prediction, *Journal of Meteorological Society of Japan* **82**(1B) (2004) 459-472.
[24] Xu G., *Theory, Algorithms and Applications*, Berlin: Springer, 2003.

Received June 20, 2013; accepted October 30, 2013.